融合文献知识聚类和复杂网络的关键技术识别方法研究*

王燕鹏[1,2]　韩涛[1,2]　陈芳[1,2]
[1] 中国科学院文献情报中心　北京 100190　[2] 中国科学院大学经济与管理学院图书情报与档案管理系　北京 100190

摘　要：[目的/意义] 立足情报研究视角，提出一套科学有效且可复用推广的关键技术识别方法，以期为国家、地区、企业和创新机构发现、部署、推动关键技术研发前瞻性布局提供情报支撑。[方法/过程] 在关键技术类型及概念界定的基础上，利用文献知识聚类识别热点技术，以各项热点技术为节点构建复杂网络，通过节点二次聚类和可视化方法展现技术结构网络，采用结构洞理论分析网络和节点特性，以此遴选共性技术；利用链路预测方法，预测技术结构网络中的缺失边产生连接的可能性，分析热点技术交叉融合促进创新技术形成的现象，以此识别潜在新兴技术。[结果/结论] 以智能制造领域为例开展关键技术识别的实证研究，通过国家权威规划文件对比和文献资料调研，初步验证方法的可操作性和有效性。

关键词：关键技术识别　复杂网络　结构洞　链路预测　技术结构
分类号：G250
DOI：10.13266/j.issn.0252-3116.2020.16.011

关键技术是指在一个系统或一项技术领域中起到重要作用且不可或缺的环节或技术，可以是当下的技术热点、难点或未来的技术突破点，也可以是对某个领域起到至关重要作用的知识。习近平总书记在 2018 年两院院士大会上提出“以关键共性技术、前沿引领技术、现代工程技术、颠覆性技术创新为突破口，敢于走前人没走过的路，努力实现关键核心技术自主可控”。在中美贸易战的背景下，我国由于技术发展不平衡而面临的一系列关键技术威胁日渐突显，例如高端光刻机、智能核心算法、锂电池隔膜、光伏逆变器等等，已对我国社会发展、经济进步、人民福祉等各方面产生了前所未有的渗透和影响[1]。在新一轮科技革命和产业变革的机遇下，清楚认识战略前沿领域的关键技术，对国家、地区和企业开展关键技术研发前瞻性布局具有重要的现实意义。鉴于此，本文立足情报研究视角开展方法探索，旨在提出一套科学有效且可复用推广的关键技术识别方法体系。

1　相关研究

实现技术识别和预见的目标，遴选可能产生最大化经济效益和社会效益的关键技术，很大程度上取决于科学合理的理论和方法体系。随着政府、学术界和产业界的深度研究实践，关键技术识别和预见方法也在不断丰富和完善。目前，常见的研究方法可归纳为如下几类：

1.1　基于专家智慧的识别方法

该类方法依靠领域专家多年研究和实践经验，结合当下社会、经济和政治发展环境，对关键技术演进方向、发展路线及措施、时间节点等进行识别和预见。具体实施方法包括德尔菲法、技术路线图法等。日本第十次科学技术预见活动[2]采用德尔菲法开展多学科、多课题调查，以识别和预见关键技术。类似的，科技部面向“十三五”科技规划编制实施技术预见调查与关键技术分析工作；中国科学院持续开展中国未来 20 年技术预见研究[3]；上海市[4]、北京市[5]也部署了关键技术识别和预见的相关工作。N. Uchihira[6]提出了制定

* 本文系中国科学院文献情报中心青年人才领域前沿项目(项目编号:2019QNGR003)研究成果之一。
作者简介：王燕鹏(ORCID:0000-0002-2583-9895)，助理研究员，硕士；韩涛(ORCID:0000-0001-5955-7813)，业务管理处处长，研究员，博士，通讯作者，E-mail:hant@mail.las.ac.cn；陈芳(ORCID:0000-0003-2517-5299)，副研究员，博士。
收稿日期:2019-12-06　修回日期:2020-03-29　本文起止页码:105-113　本文责任编辑:杜杏叶

并发系统技术路线图的方法,以识别关键技术的研发方向。

基于专家智慧的识别方法凝聚了领域专家多年的研究和实践经验,识别结果具有较强的科学性和权威性,但该类型方法受专家主观认识影响较大,新技术、新想法出现的初期,难以形成专家共识,容易遗漏部分潜在的新兴关键技术。

1.2 基于多指标评估的识别方法

该类方法通过建立多指标的评估框架来识别和预见关键技术,并结合一定的技术实例验证框架的有效性。目前,多指标评估方法大致可以分为如下三类:第一类,基于客观指标计算的定量研究方法,该类方法所采用的指标多为客观定量指标,可以从现有文献资料中采集或计算得到,取值方案多样化,如 C. Y. Lee 等[7]构建多专利指标体系并利用机器学习方法开展新兴技术的早期识别;第二类,基于主观指标评估的定性研究方法,该类方法所采用的指标多为主观评判指标,依靠专业人士对某一特定技术进行定性评判,如 D. Nagy 等[8]提出了颠覆性技术评估的三步走方法,包括技术创新性特征、价值链环节和在位技术对比。第三类,基于主观指标评分的定性与定量相结合的方法,该类方法所采用的指标多为主观评价指标,但这类指标的取值方案通常是专家打分,最终各指标分数通过统计加权计算进行汇总,以此遴选领域关键技术,因此是定性和定量相结合的方法,如中国工程院的"中国工程科技 2035 发展战略研究"[9]。

基于多指标评估的识别方法通过构建多维度指标体系对技术进行全面、系统的评估和分析,科学性和客观性较强。但同时,部分评估指标依然需要专家进行打分,对方法客观性造成一定影响。

1.3 基于模型构建的识别方法

该类方法通常是按照一定的理论框架和准则,通过统计等方式构建相关模型并对关键技术进行识别和预见,目前,常用的模型包括:TRIZ 理论、数据包络分析、突变理论、趋势外推法、样本外预测方法等。J. G. Sun 等[10]运用 TRIZ 理论构建模型,基于进化法则判断一项技术是主流技术或是相对落后的技术,并对其未来前景进行预见。李政等[11]运用数学建模法即尖点型突变方程构建关键技术的预见模型,以实现科技评价和预见。

基于模型构建的识别方法充分借鉴了经济学、运筹学和数学等学科的基础理论知识,形成了相对完整的分析框架和准则,能够更加系统客观的开展技术识别和预见活动,具有客观、简便易操作、时间周期短等特性。同时,我们也应注意到,构建模型对技术进行分析时,需要研究者具备完整的领域专业知识,充分掌握技术市场现状和趋势信息,因而该类方法对研究者的专业素养和实践技能有着较高的要求。

1.4 基于文本内容和关系的识别方法

该类方法主要基于科技文献、社交媒体等多源文本数据,利用文献计量学、文本内容和关系挖掘等手段开展关键技术识别和预见活动。目前,常用方法包括传统文献计量学方法以及融合机器学习、大数据技术和复杂网络理论等的技术识别和预见方法。如 F. Dotsika 等[12]抽取领域文献关键词并构建共现网络,通过网络关键词节点的指标特征实现新兴技术等的识别和预见;白光祖等[13]从文献知识关联的视角进行了颠覆性技术的预见;T. C. Jin 等[14]基于专利数据构建大型引用关系网络,通过分析技术主路径演化网络,发现激进式技术创新所存在的网络;周源等[15]利用 LDA 和 SVM 模型以及时间序列方法预测以技术为驱动力的新兴技术发展趋势。

基于文本内容和关系的识别方法从定量研究视角出发,所采用的数据和方法具有较强的客观性,一定程度上缓解了专家认知偏差对预见结果的影响,具有过程简洁、结果客观的特点,是开展关键技术识别和预见活动的有力工具。

综上,基于专家智慧、多指标评估和模型构建的方法在开展关键技术识别时,多需要领域专业人员的深度参与,侧重对技术的定性分析,存在识别过程成本高、主观性强、识别结果可解释性弱的问题。而基于文本内容和关系的方法侧重定量分析,通过多源文本数据和显性分析过程实现关键技术识别的目的,具有识别操作过程成本低、可重复性强、识别结果可解释性强的优势。同时,也应注意到,后者方法仍处于不断研究和尝试的阶段,尚未形成统一、认可度较高的方法体系和框架。因此,本文从文本内容和关系出发,融合机器学习方法和复杂网络理论,构建一套科学有效且可复用推广的方法体系,为关键技术识别提供新的问题解决思路。

2 融合文献知识聚类和复杂网络的关键技术识别方法

本文在关键技术类型及概念界定的基础上,利用文献知识聚类识别热点技术,以各项热点技术为节点

构建复杂网络，通过节点二次聚类和可视化方法展现技术结构网络，采用结构洞理论分析网络和节点特性，从中遴选出共性技术；利用复杂网络链路预测算法，预测技术结构网络中的缺失边产生连接的可能性，判断现有热点技术未来的交叉融合，以此识别潜在新兴技术。研究总体技术路线及方法如图 1 所示，具体来说：

2.1 关键技术类型及概念界定

在文献调研基础上，本文将关键技术定义为三类：①热点技术：某学科领域内，近年来受到科研人员广泛关注并已产出相应研究成果的主要研究方向和技术主题，反映了学科领域的研发现状和技术结构全貌；②共性技术：对学科领域内其他研究方向和技术主题产生广泛影响的，其研究成果可供参考借鉴和分享使用的一类技术，反映了学科领域的重要研发基础，是学科领域内的"思想源泉"；③潜在新兴技术：某学科领域内，未来由于科学交叉融汇而产生的新兴研究方向和技术主题，反映了学科领域内具有重要研发前景、值得深入探索的技术交叉点。

2.2 基于文献知识聚类的热点技术识别

热点技术代表了当前学科领域的主流研究方向，基本涵盖了该学科科研人员的主要研究关注点和着力点。文本聚类作为一种无监督的机器学习方法，依据同类文本相似度较大，而不同类文本相似度较小的假设，从大量文档中将内容相似的文档聚为若干类簇，结合领域专业知识开展文档判读，对每个类簇的核心科学问题和解决方案进行归纳，进而识别当前领域的热点技术，实现数学空间关系计算向领域专业知识发现的转变。基于此，本研究首先遴选出特定学科领域的核心文献数据，经过文本预处理并向量化表示，采用无监督聚类方法进行文献知识聚合，形成多个文献簇。结合人工判读和情报调研对各个文献簇进行含义解读，并将意义明确、边界清晰的文献簇作为领域热点技术。

2.2.1 文本预处理及表示

文本预处理及表示包含分词、词性还原处理、特征词提取、向量化表示过程。首先，使用 Python 的 NLTK 包对每篇文档按照空格进行分词；其次，对分词得到的逐个单词进行词性还原处理，即去除单词的词缀，提取其主干部分，如"cars→car"、"ate→eat"；然后，针对每篇文档，根据词性标注结果，选取词性为名词或动词且经过停用词过滤后的单词作为该文档的特征词，以此表征文档的主要内容；最后，使用 TF-IDF 方法对每篇文档进行向量化表示。

2.2.2 文档聚类

文档聚类使用 K-means + + 算法完成。在聚类时，需要提前确定聚类簇数量（即 K 值），一方面，计算并对比多个不同 K 值下的轮廓系数，另一方面，人工粗略评估不同 K 值下的实际聚类效果，以此综合评估聚类效果的优劣，选取最佳聚类簇数量。

2.3 基于结构洞理论的共性技术识别

共性技术反映了当前学科领域的研发基础，对该学科领域的发展起到了重要推进作用。而结构洞理论形容非冗余的联系[16]。H. J. Raider[17]的实证研究也表明：结构洞占据者的位置对信息控制、识别以及交易起着重要的作用。由此可知，结构洞反映了节点对网络资源的控制能力，占据结构洞的节点能够获取来自多方面的非重复性信息，反映在学科领域研究上，占据结构洞的研究方向更有可能是"思想源泉"，即领域共性技术，可对学科领域内其他研究方向产生广泛影响。基于此，本研究以热点技术文献簇质心向量为节点，以向量间相似性为连边依据，以此构建复杂网络，通过可视化和质心向量二次聚类的方法展现技术结构网络。基于结构洞理论，计算网络中各节点的限制度指标，以此反映节点在网络中运用结构洞的能力，节点的限制度指标值越小，代表节点拥有的结构洞越多，运用结构洞的能力越强，以此识别领域的共性技术。

2.3.1 网络构建及可视化

聚类簇质心向量是由簇中所有文档向量计算均值得到，可从空间上反映簇的中心、从逻辑上体现簇中文档的核心内容。本研究以质心向量代表各项热点技术作为网络节点，使用余弦相似度算法计算两两质心向量间的相似性，作为网络节点连边及其权重赋予的依据，以此构建网络。采用 Gephi 软件对该网络进行可视化，并利用软件集成的模块化算法[18]对质心向量进行二次聚类，将热点技术划分为若干技术大类。

2.3.2 结构洞计算

以两两质心向量间的余弦相似度构建对称矩阵，输入到 Ucinet 软件中，通过"Network → Ego Networks → Structural Holes"计算网络中各个节点的限制度指标值。

2.4 基于复杂网络链路预测的潜在新兴技术识别

跨学科领域的交叉融合和集成创新已成为当代科学技术创新的鲜明特征，通过预测多个研究方向和技术主题产生交叉融合的可能性，将有助于识别潜在新兴技术。基于此，本研究以上述构建的技术结构网络为基础，利用复杂网络链路预测方法，探测技术结构网

络中尚未产生连接但未来可能产生较强连接的节点对，分析热点技术交叉融合促进科学技术创新形成的现象(即潜在的科学发现)，以此识别潜在新兴技术。

具体的，本研究选用链路预测中基于网络结构相似性的方法，该方法计算复杂度低、准确性好，并且网络的拓扑结构性质能够帮助选择合适的相似度指标。由于上述技术结构网络为加权无向网络，选择使用基于网络结构相似性的含权的 Adiamic-Adar 指标(简称 AA 指标)计算产生连接的可能性，以综合考虑节点共同邻居数量及节点度对连接的影响，使链路预测结果更加准确。本文研究技术路线图如图 1 所示：

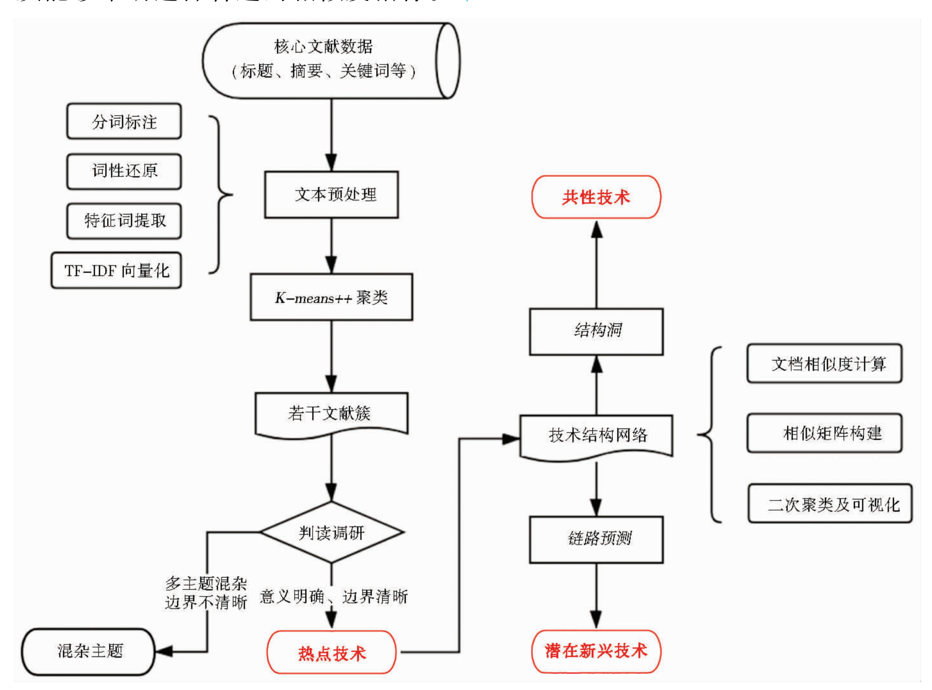


图 1　研究技术路线图

3　关键技术识别分析实证

为检验本文融合文献知识聚类和复杂网络的关键技术识别方法的可操作性和有效性，以“智能制造”领域为例开展实证研究。智能制造是伴随信息技术的不断普及而逐步发展起来的，是基于新一代信息通信技术与先进制造技术深度融合，贯穿于设计、生产、管理、服务等制造活动的各个环节，具有自感知、自学习、自决策、自执行、自适应等功能的新型生产方式。

3.1　核心文献数据遴选

本文以科技论文为分析对象，依据 *Journal Cited Report* 期刊引证报告，选取“工程、制造”领域下的 Q1 区和 Q2 区期刊，以及“工程，电气与电子”“工程，机械”“工程，交叉学科”“计算机科学，人工智能”“计算机科学，控制论”“计算机科学，跨学科应用”领域下的 Q1 区期刊。根据期刊名称、简介和典型论文判别各期刊是否与智能制造领域有交叉和相关性，共遴选出 26 种期刊(见表 1)，并将其中 7 种期刊作为密切相关期刊，另外 19 种为相关期刊。在构建智能制造知识体系的基础上，制定检索词。以密切相关期刊全部载文、相关期刊载文限定检索词的方式作为检索策略，在 Web of Science 核心合集中检索论文，时间跨度为 2015 – 2018 年，共 15 345 篇论文，作为关键技术识别的核心文献数据。

3.2　热点技术识别分析及技术结构网络构建

以上述核心文献的标题、摘要、关键词字段内容作为分析数据，进行文本预处理、向量化表示处理、K-means + + 聚类，结合轮廓系数的计算结果，确定聚类文献簇数量 K = 100 时，具有较好的聚类效果。经人工判读，100 个聚类文献簇中共有 98 个有效文献簇和 2 个混杂文献簇，本文以 98 个有效文献簇作为智能制造领域的热点技术。以 98 个热点技术作为节点构建技术结构网络，并进行可视化和二次聚类划分，可形成 10 项技术大类，见图 2 及表 2。

在智能制造技术方面，共包含 3 项技术大类：机器人及工业控制技术、人工智能技术、物联网及传感器技术。其中，热点技术主要包括工业机器人、机器人跟踪自适应、电机设计与控制、执行器与驱动器等机器人核

表 1　智能制造科技期刊遴选情况(共 26 种)

分类	期刊分区	期刊名称	相关度
工程,制造	Q1	*international Journal of Machine Tools & Manufacture*	相关
工程,制造	Q1	*Composites Part A-Applied Science and Manufacturing*	相关
工程,制造	Q1	*Ieee-Asme Transactions on Mechatronics*	密切相关
工程,制造	Q1	*international Journal of Precision Engineering and Manufacturing-Green Technology*	密切相关
工程,制造	Q1	*Journal of Manufacturing Systems*	相关
工程,制造	Q1	*Journal of intelligent Manufacturing*	密切相关
工程,制造	Q1	*Cirp Annals-Manufacturing Technology*	相关
工程,制造	Q1	*Journal of Manufacturing Processes*	相关
工程,制造	Q2	*international Journal of Advanced Manufacturing Technology*	密切相关
工程,制造	Q2	*Flexible Services and Manufacturing Journal*	密切相关
工程,制造	Q2	*3D Printing and Additive Manufacturing*	密切相关
工程,制造	Q2	*international Journal of Computer integrated Manufacturing*	密切相关
工程,电气与电子	Q1	*Ieee Transactions on Pattern Analysis and Machine intelligence*	相关
工程,电气与电子	Q1	*Ieee Communications Magazine*	相关
工程,电气与电子	Q1	*Ieee Wireless Communications*	相关
工程,电气与电子	Q1	*Automatica*	相关
工程,电气与电子	Q1	*Ieee internet of Things Journal*	相关
工程,电气与电子	Q1	*Ieee Transactions on Automatic Control*	相关
工程,工业	Q1	*Ieee Transactions on industrial informatics*	相关
工程,交叉学科	Q1	*Archives of Computational Methods in Engineering*	相关
工程,交叉学科	Q1	*integrated Computer-Aided Engineering*	相关
工程,交叉学科	Q1	*Advanced Engineering informatics*	相关
工程,交叉学科	Q1	*Advances in Engineering Software*	相关
工程,交叉学科	Q1	*Engineering Applications of Artificial intelligence*	相关
计算机科学,控制论	Q1	*Ieee Transactions on Cybernetics*	相关
计算机科学,跨学科应用	Q1	*Computers & industrial Engineering*	相关

心技术;自适应控制、车辆动力学与控制、模型预测控制、滑膜控制、延迟控制、一致性控制等工业控制技术;机器学习、模式挖掘、计算机视觉等人工智能方法在制造过程中的应用。此外,还包括传感器、物联网计算、物联网安全、网络制造等物联网核心技术研究及应用。

在制造加工工艺方面,共包含 4 项技术大类:成型工艺、铣削工艺、焊接工艺、特种加工工艺。其中,热点技术主要包括新材料性能及 3D 打印、激光选区融化与激光加工技术等增材制造核心技术环节;注射成型、挤压成型、锻造成型、液压成型、深拉伸成型等成型工艺研究及优化;刀具路径、刀具磨损、误差估计、颤振稳定性、切削力建模优化、切屑形成等铣削基础工艺原理的研究及优化;热处理加工、应力分析、加工变形、搅拌摩擦等焊接工艺优化,以及电火花加工、粉末冶金等特种加工工艺。

在智能制造工艺方面,共包含 3 项技术大类:先进制造模式及设计优化、工艺规划与生产调度、制造过程监测与控制。其中,热点技术主要包括可制造性设计理论、单元式制造、增材制造、再制造等先进制造模式探索;产品设计与优化方法、库存管理、设施布局优化、供应链优化等制造过程的设计优化研究;智能装配、生产计划与调度控制、项目管理等工艺规划与调度研究;制造系统可靠性评估、分布式状态估计、控制图优化、预测性维护、故障监测与隔离等制造过程监测与控制研究。

3.3　共性技术识别分析

以智能制造领域热点技术结构网络(见图 2)为对象,计算网络中各节点的限制度指标并升序排列,取 Top20 的热点技术作为智能制造领域的共性技术(见表 3)。

在智能制造工艺方面,共包含 11 项共性技术,例如制造过程监测与控制论及方法、可制造性设计理论、先进制造模式探索、单元式制造、增材制造等。

在制造加工工艺方面,共包含 6 项共性技术,例如精密机床、新材料性能及 3D 打印、激光选区融化与激

光加工技术、注射成型工艺、刀具路径规划等。

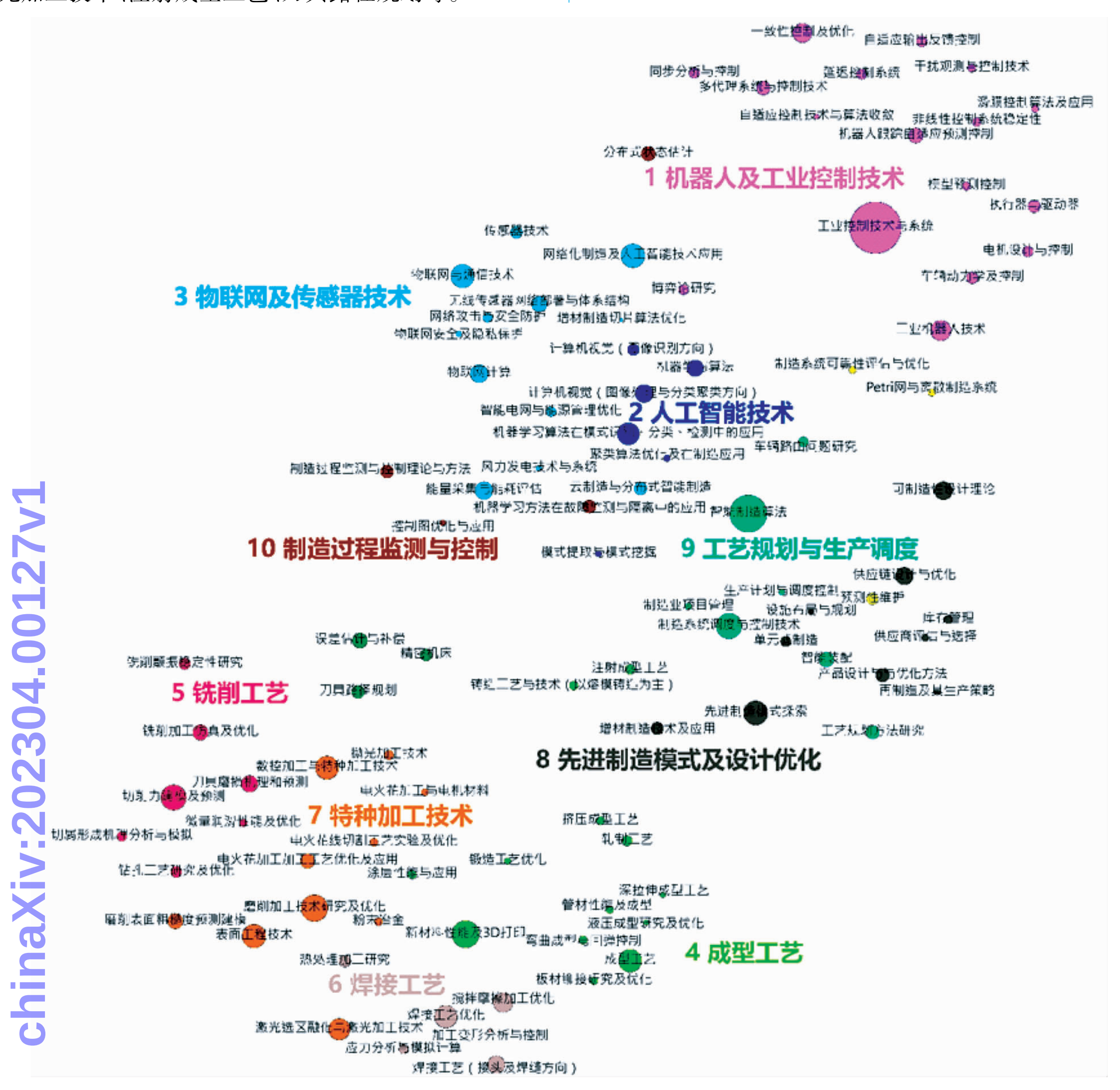


图 2 智能制造领域技术结构网络

注:图中每个节点代表一个聚类文献簇,即热点技术,节点大小代表聚类簇下文章数量多少,节点颜色代表该节点归属的技术大类,“数字 - 标签”代表同颜色热点技术所归属的技术大类

表 2 智能制造领域热点技术

序号	技术大类	热点技术
1	机器人及工业控制技术	工业机器人技术 ｜ 电机设计与控制 ｜ 执行器与驱动器 ｜ 自适应控制技术与算法收敛 ｜ 工业控制技术与系统 ｜ 车辆动力学及控制 ｜ 多代理系统与控制技术 ｜ 机器人跟踪自适应预测控制 ｜ 模型预测控制 ｜ 滑膜控制算法及应用 ｜ 延迟控制系统 ｜ 博弈论研究 ｜ 非线性控制系统稳定性 ｜ 干扰观测与控制技术 ｜ 自适应输出反馈控制 ｜ 同步分析与控制 ｜ 一致性控制及优化
2	人工智能技术	机器学习算法在模式识别、分类、检测中的应用 ｜ 机器学习算法 ｜ 模式提取与模式挖掘 ｜ 计算机视觉(图像识别方向) ｜ 计算机视觉(图像处理与分类聚类方向) ｜ 聚类算法优化及在制造中的应用
3	物联网及传感器技术	能量采集与能耗评估 ｜ 网络化制造及人工智能技术应用 ｜ 智能电网与能源管理优化 ｜ 传感器技术 ｜ 风力发电技术与系统 ｜ 物联网与通信技术 ｜ 增材制造切片算法优化 ｜ 云制造与分布式智能制造 ｜ 无线传感器网络部署与体系结构 ｜ 网络攻击与安全防护 ｜ 物联网计算 ｜ 物联网安全及隐私保护
4	成型工艺	新材料性能及 3D 打印 ｜ 注射成型工艺 ｜ 铸造工艺与技术(以熔模铸造为主) ｜ 挤压成型工艺 ｜ 成型工艺 ｜ 轧制工艺 ｜ 锻造工艺优化 ｜ 板材铆接研究及优化 ｜ 弯曲成型与回弹控制 ｜ 液压成型研究及优化 ｜ 深拉伸成型工艺 ｜ 管材性能及成型 ｜ 涂层性能与应用

（续表2）

序号	技术大类	热点技术
5	铣削工艺	精密机床 丨 刀具路径规划 丨 误差估计与补偿 丨 铣削颤振稳定性研究 丨 铣削加工仿真及优化 丨 钻孔工艺研究及优化 丨 刀具磨损机理和预测 丨 切削力建模及预测 丨 微量润滑性能及优化 丨 切屑形成机理分析与模拟
6	焊接工艺	热处理加工研究 丨 应力分析与模拟计算 丨 加工变形分析与控制 丨 焊接工艺优化 丨 搅拌摩擦加工优化 丨 焊接工艺(接头及焊缝方向)
7	特种加工技术	激光选区融化与激光加工技术 丨 数控加工与特种加工技术 丨 电火花加工与电机材料 丨 表面工程技术 丨 粉末冶金 丨 电火花加工工艺优化及应用 丨 磨削表面粗糙度预测建模 丨 电火花线切割工艺实验及优化 丨 抛光加工技术 丨 磨削加工技术研究及优化
8	先进制造模式及设计优化	可制造性设计理论 丨 先进制造模式探索 丨 单元式制造 丨 产品设计与优化方法 丨 增材制造技术及应用 丨 库存管理 丨 设施布局与规划 丨 供应链设计与优化 丨 再制造及其生产策略 丨 供应商评估与选择
9	工艺规划与生产调度	工艺规划方法研究 丨 智能制造算法 丨 智能装配 丨 生产计划与调度控制 丨 制造系统调度与控制技术 丨 制造业项目管理 丨 车辆路由问题研究
10	制造过程监测与控制	制造过程监测与控制的理论与方法 丨 制造系统可靠性评估与优化 丨 分布式状态估计 丨 Petri 网与离散制造系统 丨 控制图优化与应用 丨 预测性维护 丨 机器学习方法在故障监测与隔离中的应用

表3　智能制造领域共性技术

共性技术	类簇内文献数量	限制度(对数值)	技术研究内容示例
制造过程监测与控制理论及方法	129	－3.09	过程智能监控、制造设备状态监控、损伤检测、遗传算法、混合马尔科夫模型、多传感器目标监测、传感器融合监测
可制造性设计理论	190	－3.08	多目标优化设计、产品设计优化集成、计算机辅助夹具设计、基于决策的设计、设计概念选择、均匀设计
先进制造模式探索	262	－3.062	智能云制造、混合集成制造、混合制造与智能工厂、矩阵结构制造系统、动态制造网络、制造过程可持续性性能分析
智能制造算法	415	－3.061	新进化算法、分布式优化神经动力学方法、退化扩散随机系统、遗传算法、改进局部搜索
工艺规划方法研究	135	－3.055	随机规划、运动规划、柔性工艺规划、集成工艺规划、启发式算法、增强拆卸规划、可持续制造工业规划
制造系统可靠性评估与优化	66	－3.055	规范粒子群优化算法、可靠性评估、多目标可靠性优化、可靠性仿真、失效模型与效应分析、退化建模、冗余分配
能量采集与能耗评估	183	－3.043	能耗特性、能量建模、能量采集器、能量传输信道模型、能量属性继承、能源管理
精密机床	115	－3.03	五轴数控机床、精密机床、螺旋锥齿轮数控机床加工理论、超精密机床、多轴数控机床、大容量机床
新材料性能及3D打印	297	－3.026	裂纹碳化物、镁合金、铁基合金、夹层材料、微米铜线、铝合金、RB-SiC 陶瓷、3D 打印、熔融沉积成型、激光融化
注射成型工艺	94	－3.024	超声注射成型、微注射成型、自动化注塑工艺、动态注塑成型、注射成型优化研究
单元式制造	100	－3.021	单元制造系统、动态单元构建、成组技术、多选择目标规划、多工艺路线、遗传算法、分组效能指数
刀具路径规划	129	－3.016	刀具轨迹、刀具半径补偿、五轴刀具轨迹生成、刀具姿态、刀具路径、刀具干涉、
产品设计与优化方法	139	－3.012	模块化产品设计、产品形态优化、低成本产品设计、动态设计与再设计、缺陷检测、产品可持续性评价、集成设计、复杂产品适应性设计
智能装配	152	－3.008	自动装配、装配线平衡、移动装配、深度成像传感器、装配线工人分配与平衡、加权装配
分布式状态估计	144	－3.006	状态估计、非渐进估计、水下机器人方向估计、分布式状态估计、人体姿态估计、卡尔曼滤波、传感器网络、自适应估计
激光选区融化与激光加工技术	230	－3.005	激光选区融化、全局灵敏度分析、微抛光、激光工程网成型、激光钎焊工艺、激光束加工、激光辅助铣削、自适应控制
增材制造技术及应用	124	－3.005	纤维增强复合材料、选择性抑制烧结、基于分布式3D打印服务的定制生产、非装配式可控气动机器人3D打印、4D打印、增材制造参数多目标优化、分层回转体结构3D打印
热处理加工研究	115	－3.005	热循环、热流密度、磁流变、快速热烧蚀、热机械耦合、有效热容、热模拟、热建模、温度场影响、材料热导率、热分析
智能电网与能源管理优化	91	－3.004	智能电网、最优能源管理、电动汽车预测储能功率管理随机控制、充电遗传模糊调度、在线能源管理优化
电机设计与控制	105	－3.001	永磁同步电机、纳米步进电机、气动电机控制、无人机电机异常检测、直线转矩控制、制动力矩控制、粒子群优化

在智能制造技术方面，共包含3项共性技术，例如能量采集与能耗评估、智能电网与能源管理优化、电机设计与控制。

为验证上述识别结果的准确性，选取工业和信息化部、财政部在2016年12月8日印发的《智能制造发展规划(2016－2020年)》(以下简称"规划")为对照依据，以规划中提到的12项智能制造关键共性技术创新方向对比上述识别出的20项共性技术。从对比结果(见表4)可以看出，除高可靠实时通信外，识别出的20项共性技术基本可以涵盖规划中的12项关键共性技术创新方向，且技术层级更低、颗粒度更细、技术范畴更明确。

表4 规划中智能制造关键共性技术创新方向与识别结果的对比

规划中提到的智能制造关键共性技术创新方向	本研究识别出的智能制造共性技术
新型传感技术	制造过程监测与控制理论及方法(多传感器目标监测、传感器融合监测、过程智能监控) 智能装配(深度成像传感器)
模块化/嵌入式控制系统设计技术	可制造性设计理论(多目标协同优化设计、产品设计优化集成)
先进控制与优化技术	制造过程监测与控制理论及方法(多传感器目标监测、传感器融合监测、过程智能监控)
系统协同技术	工艺规划方法研究(集成工艺规划) 单元式制造(成组技术)
故障诊断与健康维护技术	制造过程监测与控制理论及方法(制造设备状态监控、损伤检测)
高可靠实时通信	——
功能安全技术	制造过程监测与控制理论及方法(传感器融合监测、过程智能监控)
特种工艺与精密制造技术	精密机床、增材制造、超声注射成型、激光钎焊工艺、裂纹碳化物
识别技术	智能装配(自动装配、装配线平衡、移动装配、深度成像传感器)
建模与仿真技术	计算机辅助夹具设计、可靠性仿真、失效模型
工业互联网	智能电网
人工智能	智能制造算法、制造系统可靠性评估与优化、智能装配

3.4 潜在新兴技术识别分析

以智能制造领域热点技术结构网络(见图2)为基础,预测网络中的缺失边产生连接的可能性,取Top10的热点技术节点对作为智能制造领域的潜在新兴技术,如表5所示:

表5 智能制造领域潜在新兴技术

潜在新兴技术		产生连接的可能性
新材料性能及3D打印	可制造性设计理论	0.192
液压成型研究及优化	新材料性能及3D打印	0.083
管材性能及成型	板材铆接研究及优化	0.067
网络化制造及人工智能技术应用	先进制造模式探索	0.056
机器人跟踪自适应预测控制	可制造性设计理论	0.054
制造系统调度与控制技术	制造系统可靠性评估与优化	0.052
机器学习算法在模式识别、分类、检测中的应用	工业控制技术与系统	0.051
无线传感器网络部署与体系结构	工业控制技术与系统	0.048
电火花加工工艺优化及应用	切削力建模及预测	0.043
非线性控制系统稳定性	网络化制造及人工智能技术应用	0.042

在智能制造领域的热点技术中,可制造性设计理论在新材料性能与3D打印中的应用、新材料性能与3D打印和液压成型工艺相结合可能会产生深度耦合和交叉,成为未来的新兴技术。此外,网络化制造及人工智能技术应用对先进制造模式的探索突破、机器学习算法和无线传感器网络在工业控制中的应用等也是智能制造领域的潜在新兴技术。

为检验上述识别结果的合理性,以"可制造性设计理论在新材料性能与3D打印中的应用"为例开展文献资料调研论证。大连理工大学工业装备结构分析国家重点实验室刘树田教授[19]提出,现有的增材制造结构绝大部分仍然采用面向传统制造工艺的设计构型,未充分利用增材制造所提供的新型设计空间,性能无法在本质上得到飞跃,且受限于增材制造技术的不成熟,其性能甚至劣于传统制造工艺所制备的结构。因此,开展面向增材制造的设计方法研究,发展完整的设计理论体系,突破传统设计极限,获得优质创新结构构型是非常重要的研究方向。由此文献案例可以看出,"可制造性设计理论在新材料性能与3D打印中的应用"作为潜在新兴技术具有一定的合理性。

4 结语

近年来,随着"数据驱动"的观念深化,机器学习、大数据技术和复杂网络理论等的逐步普及为关键技术识别带来了新的问题解决思路。本文立足情报研究视角开展方法探索,以关键技术识别为应用目的,以科技创新基本规律为理论基础,以文献知识聚类和技术结构复杂网络构建为分析主线,以无监督聚类、结构洞理论和复杂网络链路预测为分析方法,进行热点技术、共性技术和潜在新兴技术的挖掘和识别,并以智能制造领域为例开展实证研究,通过国家权威文件对比和文献资料调研,初步验证本文方法的可操作性和有效性。

囿于研究能力有限,本文所提出的方法在以下方面仍有改进空间:首先,文本预处理过程中,词形还原的准确性有待进一步提高;其次,将结构洞理论与其它网络及客观度量指标相结合,多维度共同揭示共性技术;最后,聚类文献簇的解读需要领域专家或情报人员的参与,仍存在一定的主观性。后续研究将对上述问题进行改进和完善。

参考文献:

[1] 汪雪锋,赖院根,朱东华. 技术威胁理论研究[J]. 科学学研究, 2009,27(2):166-169.
[2] 七丈直弘, 小笠原敦. 第10回科学技術予測調査(ビジョン): 国際的視点からのシナリオプランニング[J]. 年次学術大会講演要旨集, 2015, 30(1): 882-885.
[3] 中国科学院技术预见研究组. 中国未来20年技术预见[M]. 北京: 科学出版社, 2006.
[4] 李红, 孙绍荣, 刘继云. 上海市科技发展重点领域技术预见的

实证研究[J]. 科学学研究, 2005, 23(s1):101－105.

[5] 穆荣平. 北京技术预见:实践与思考[J]. 世界科学, 2003, 15(4):41－43.

[6] UCHIHIRA N. Future direction and roadmap of concurrent system technology[J]. Ieice transactions on fundamentals, 2007, 90(11):2443－2448.

[7] LEE C, KWON O, KIM M, et al. Early identification of emerging technologies: a machine learning approach using multiple patent indicators[J]. Technological forecasting and social change, 2018, 50(127): 291－303.

[8] NAGY D, SCHUESSLER J, DUBINSKY A. Defining and identifying disruptive innovations[J]. Industrial marketing management, 2016, 46(57):119－126.

[9] 王昆声, 周晓纪, 龚旭,等. 中国工程科技 2035 技术预见研究[J]. 中国工程科学, 2017, 19(1):34－42.

[10] SUN J, GAO J, YANG B, et al. Achieving disruptive innovation - forecasting potential technologies based upon technical system evolution by TRIZ[C]// IEEE international conference on management of innovation and technology. Bangkok: IEEE, 2008:18－22.

[11] 李政, 罗晖, 李正风,等. 基于突变理论的科技评价方法初探[J]. 科研管理, 2017,38(s1):193－200.

[12] DOTSIKA F, WATKINS A. Identifying potentially disruptive trends by means of keyword network analysis[J]. Technological forecasting & social change,2017,49(119):114－127.

[13] 白光祖,郑玉荣,吴新年,等. 基于文献知识关联的颠覆性技术预见方法研究与实证[J]. 情报杂志,2017,36(9):38－44.

[14] JIN T, MIYAZAKI K, KAJIKAWA Y. Identification of evolutionary characteristics of emerging technologies: the case of smart grid in Japan[C]// Portland international conference on management of engineering and technology. Honolulu, HI: IEEE, 2017:649－656.

[15] 董放, 董放, 刘宇飞, 等. 基于 LDA-SVM 论文摘要多分类新兴技术预测[J]. 情报杂志, 2017, 36(7):40－45.

[16] 刘军. 整体网分析讲义: UCINET 软件实用指南[M]. 上海:格致出版社/上海人民出版社, 2009.

[17] RAIDER H J. Market structure and innovation[J]. Social science research, 1998, 27(1): 1－21.

[18] BLONDEL V D, GUILLAUME J L, LAMBIOTTE R, et al. Fast unfolding of communities in large networks[J]. Journal of statistical mechanics theory & experiment, 2008:P10008.

[19] 刘书田, 李取浩, 陈文炯, 等. 拓扑优化与增材制造结合: 一种设计与制造一体化方法[J]. 航空制造技术, 2017, 60(10): 26－31.

作者贡献说明:

王燕鹏:研究思路设计、实验设计及结果分析、论文撰写;

韩涛:研究思路修改、论文修改;

陈芳:研究思路讨论、实验结果分析。

Identification of Key Technologies Based on Literature Knowledge Clustering and Complex Network

Wang Yanpeng[1,2] Han Tao[1,2] Chen Fang[1,2]

[1] National Science Library, Chinese Academy of Sciences, Beijing 100190

[2] Department of Library, Information and Archives Management, School of Economics and Management, University of Chinese Academic of Sciences, Beijing 100190

Abstract: [**Purpose/significance**] This paper try to propose a scientific, effective and reusable method to identify key technologies based on the perspective of intelligence research. It aims to provide information support for nation, regions, enterprises and innovative institutions to discover, deploy and promote the prospective R&D of key technologies. [**Method/process**] Based on the definition of key technology and its types, this paper used K-means++ algorithm to cluster scientific papers to identify hotspot technologies. Then it used the hotspot technologies as nodes to construct and visualize complex network through secondary clustering and Gephi. Structural holes theory was adopted to analysis the network and attributes of nodes, and thereby selected generic technologies. Link prediction algorithm was used to predict the missing edges in the network according to the structure, and we can identify the potential emerging technologies based on the phenomenon of cross-fusion of hot technologies to promote the formation of innovative technologies . [**Result/conclusion**] Taking the Intelligent Manufacturing as an example to carry out empirical research on the method, and validated the operability and effectiveness of the method through national authoritative documents and literature research.

Keywords: key technology identification complex network structural holes link prediction technical structure